NADA R. SANDERS, PH.D.
Distinguished Professor of Supply Chain Management,
D'Amore-McKim School of Business, Northeastern University

BIG DATA DRIVEN SUPPLY CHAIN MANAGEMENT

A Framework for Implementing Analytics and Turning Information into Intelligence

Big Data Driven Supply Chain Management

A Framework for Implementing Analytics and Turning Information into Intelligence

Nada R. Sanders, Ph.D.

Distinguished Professor of Supply Chain Management
D'Amore-McKim School of Business
Northeastern University

Associate Publisher: Amy Neidlinger
Executive Editor: Jeanne Glasser Levine
Development Editor: Natasha Torres
Operations Specialist: Jodi Kemper
Cover Designer: Alan Clements
Managing Editor: Kristy Hart
Senior Project Editor: Lori Lyons
Copy Editor: Karen Annett
Proofreader: Debbie Williams
Indexer: Erika Millen
Senior Compositor: Gloria Schurick
Manufacturing Buyer: Dan Uhrig

Pearson Education, Inc.
Upper Saddle River, New Jersey 07458

For information about buying this title in bulk quantities, or for special sales opportunities (which may include electronic versions; custom cover designs; and content particular to your business, training goals, marketing focus, or branding interests), please contact our corporate sales department at corpsales@pearsoned.com or (800) 382-3419.

For government sales inquiries, please contact governmentsales@pearsoned.com.

For questions about sales outside the U.S., please contact international@pearsoned.com.

Printed in the United States of America

First Printing June 2014

ISBN-10: 0-13-380128-4
ISBN-13: 978-0-13-380128-6

Pearson Education LTD.
Pearson Education Australia PTY, Limited.
Pearson Education Singapore, Pte. Ltd.
Pearson Education Asia, Ltd.
Pearson Education Canada, Ltd.
Pearson Educación de Mexico, S.A. de C.V.
Pearson Education—Japan
Pearson Education Malaysia, Pte. Ltd.

Library of Congress Control Number: 2014934904

This book is dedicated to the reader.
With knowledge and understanding,
we can make our enterprising efforts more
efficient, effective, and intelligent.

This page intentionally left blank

Contents

Table of Contents

Foreword

The surge of professional and academic interest in the topic of big data analytics has been similar to the gold rush—there is much commotion but few know where to look for the real payoff. Although it has long been obvious that utilizing emerging information technology is vital to remaining competitive, the nature of that technology has shifted in a tectonic way, and thought leadership must keep pace.

Many companies have yet to leverage big data analytics to transform their supply chain operations. Many are awash in data but are unsure how to use it to drive their supply chains. Many are engaging in fragmented utilization or implementation rather than a systematic and coordinated effort. The results are isolated benefits, lack of insight and competitiveness, and supply chains plagued with inefficiencies and cost overruns. Others are unsure how to even begin, particularly small and medium-sized firms.

Why all the confusion? The reason is that companies lack a clear roadmap for how to implement big data analytics in a meaningful and cost-effective manner.

This book attempts to remedy the situation by providing a systematic framework for companies on *how* to implement big data analytics across the supply chain to turn information into intelligence and achieve a competitive advantage. This end-to-end perspective on the application of big data analytics provides a much-needed conceptual organization to this topic while linking strategy with tactics. Furthermore, this roadmap shows organizational leaders *how* to implement the type of organizational change big data analytics requires.

Acknowledgments

I am most grateful to my family—especially my parents, who taught me how to create order out of chaos, and convey with clarity.

I would also like to thank all the professionals who shared their experiences and provided insights into this work. Special thanks goes to Jeanne Glasser Levine at Pearson, as well as the other talented staff at Pearson who have contributed work to this book.

About the Author

Nada R. Sanders, Ph.D., is Distinguished Professor of Supply Chain Management at the D'Amore-McKim School of Business at Northeastern University, and she holds a Ph.D. from the Ohio State University. She is an internationally recognized thought leader and expert in forecasting and supply chain management. She is author of the book *Supply Chain Management: A Global Perspective* and is coauthor of the book *Operations Management,* in its 5th edition. She was ranked in the top 8 percent of individuals in the field of operations management from a pool of 738 authors and 237 different schools by a study of research productivity in U.S. business schools. Nada is a Fellow of the Decision Sciences Institute, and has served on the Board of Directors of the International Institute of Forecasters (IIF), Decision Sciences Institute (DSI), and the Production Operations Management Society (POMS). Her research focuses on the most effective ways for organizations to use technology to achieve a competitive advantage.

This page intentionally left blank

1

A Game Changer

The era of radically different competition is here. It is a tsunami that has transformed entire industries and left numerous casualties in its wake. Like Gutenberg's invention of the printing press changing the world through printing, the move toward big data is creating an equally tectonic shift in business and society. Transform or be left behind.

Consider the fate of Borders.[1] In l971, the company opened its first store in Ann Arbor, Michigan, when the book industry was a different place. In 2011, 40 years later, the bookstore chain closed its doors. So, what happened? Borders fell behind the curve on embracing the Web and the digital world of data. Not understanding that the rules of the game had changed, Borders had outsourced its online bookselling to Amazon.com. So any time you visited Borders.com, you were redirected to Amazon. Playing by the old rules made this seem like a smart decision. In the new world, however, there was a problem.

To jump on the tails of Amazon and leverage its competitive priorities did not take into account that playing in the digital world was *the* competitive priority. Relinquishing control to another company would simply cut into the company's customer base. Also, not understanding that the world was now a digital place, Borders did not embrace e-books, like Amazon and Barnes & Noble. Walking into Borders was like walking into a bookstore of yesteryear. The outcome was predictable.

The competitive world Borders lived in was one where booksellers tracked which books sold and which did not. Loyalty programs could help tie purchases to individual customers. That was about it. Then shopping moved online. The ability to understand and track customers changed dramatically. Online retailers could track every aspect of what customers bought. They could track what customers looked at, how they navigated through the site, how long they hovered over a site, and how they were influenced by promotions and page layouts. They were now able to develop microsegments of individual customers and groups based on endless characteristics. They could then create individually targeted promotions. Then algorithms were developed to predict what books individual customers would like to read next. These algorithms were self-teaching and performed better every time the customer responded to a recommendation. Traditional retailers like Borders simply couldn't access this kind of information. They could not compete in a timely manner.

And Amazon? With its Kindle e-book readers and convincing hundreds of publishers to release their books on the Kindle format, the company has cornered the market. It has "datafied" books—turning them into quantified format that can be tabulated and analyzed.[2] This allows Amazon everything from recommending books to using algorithms to find links among the topics of books that might not otherwise be apparent. Embracing the digital age, technology, and data-driven decisions, the company is moving well beyond wanting to be the biggest bookstore on the Internet. It is moving toward being the most dominant retailer in the world. Amazon understands that this means using big data and technology to manage its entire supply chain in a synchronized manner. In fact, Jeff Bezos, Amazon's CEO, is known for demanding rigorous quantification of customer reactions before rolling out new features.[3] Data and technology have been used to coordinate everything from customer orders to fulfillment, inventory management, labor, warehousing, transportation, and delivery.

Amazon is not the only one. Leading-edge companies across the globe have scored successes in their use of big data. Consider Walmart, Zara, UPS, Tesco, Harrah's, Progressive Insurance, Capital One, Google, and eBay.[4] These companies have succeeded in this game-changing environment by embracing and leading the change. They have used big data analytics to extract new insights and create new forms of value in ways that have changed markets, organizations, and business relationships.

1.1 Big Data Basics

To fully understand the impact of big data analytics, we first need to have a clear idea of what it actually is. In this section we explain big data basics. We define the key concepts of big data analytics and explain how these concepts find novel applications across business. This will set up the book's subsequent discussions of big data analytics applications in supply chain management.

1.1.1 Big Data

Big data is simply lots of data. The term *big data* specifically refers to large data sets whose size is so large that the quantity can no longer fit into the memory that computers use for processing. This data can be captured, stored, communicated, aggregated, and analyzed.[5] There is no specific definition of the size of big data, such as the number of terabytes or gigabytes. The reason is that this is a moving target. Technology is advancing over time and the size of data sets that are considered big data will also increase.

As the volume of data has grown, so has the need to revamp the tools used for analyzing it. That is how new processing technologies like Google's MapReduce and its open source equivalent, Hadoop, were developed. These new technologies enable companies to manage far-larger quantities of data than before. Most important, unlike in the past, this data does not need to be placed in neat rows and columns as traditional data sets to be analyzed by today's technology.

Big data comes in different forms. It includes all kinds of data from every source imaginable. It can be structured or unstructured. It can be a numerical sequence or voice and text and conversation. It can come in the form of point-of-sale (POS), radio-frequency identification (RFID), or Global Positioning System (GPS) data, or it can be in the form of Twitter feeds, Facebook, call centers, or consumer blogs. Today's advanced analytical tools allow us to extract meaning from all types of data.

1.1.2 Analytics

Analytics is applying math and statistics to these large quantities of data. When we apply math and statistics to big data—often called *big data analytics*—we can gain insights into the world around us unlike ever before. We can infer probabilities or likelihoods that something will happen.

We are used to this in our everyday life. We are accustomed to e-mail filters that estimate the likelihood that an e-mail message is spam or that the typed letters *teh* are supposed to be *the*. The key is that these systems perform well because they are fed with lots of data on which to base their predictions. Moreover, the systems are built to improve themselves over time, by keeping tabs on the best signals and patterns to look for as more data is fed in. Think about "teaching" your e-mail filter that a type of e-mail is a spam by labeling similar e-mails.

It is through big data that Walmart learned that customers prefer to stock up on the sugary treat Pop-Tarts during a hurricane,[6] eBay identified which Web designs generate the highest sales,[7] and Progressive Insurance learned how to optimize insurance premiums by risk category.[8]

Even small companies have benefited. Consider the online music equipment retailer The Musician's Friend. Using basic analytics, the company was able to compare different versions of its Web page to identify customer preferences. The preferred site generated a 35 percent increase in sales over the original home page. This simple change resulted in a measurable improvement on return on investment (ROI).[9]

1.1.3 Big Data and Analytics: The Perfect Duo

To set the record straight, big data without analytics is just lots of data. We've been accumulating a lot of data for years. Analytics without big data is simply mathematical and statistical tools and applications. Tools such as correlation and regression, for example, have been around for decades. In fact, Google's director of research, Peter Norvig, explained it well by saying: "We don't have better algorithms. We just have more data."[10]

However, it is the combination that makes the difference. It is through the combination of big data and analytics that we can get the really meaningful insights and turn information into business intelligence (see Figure 1.1). Also, big data and analytics build on each other. Continued application of even simple analytical tools results in their improvement, refinement, and sophistication. Consider that as you increasingly identify the number of e-mails as spam, the filter "learns" and becomes better at correctly identifying spam. It is for this reason we use the term *big data analytics* throughout this book to refer to the application of analytics to these large data sets.

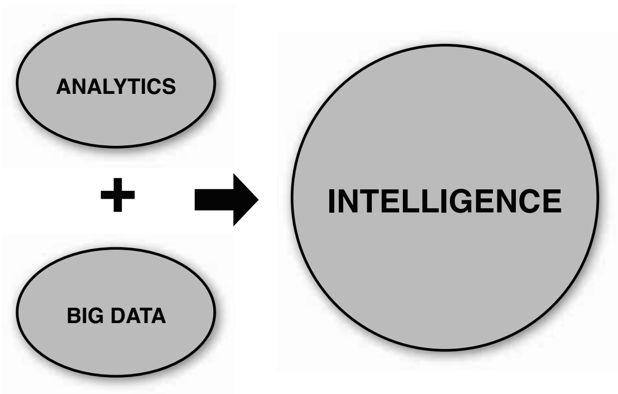

Figure 1.1 Turning information into intelligence

1.1.4 New Computing Power

How can companies extract intelligence out of these huge amounts of data? This is made possible through today's massive computing power available at a lower cost than ever before. Large data, coupled with larger and more affordable computing power, means that you can do on a larger scale that which cannot be done on a smaller one. Improvements in computing have resulted in large advances in capability. This has enabled high-level analytics to be performed on these large and unstructured data sets.

Processing power has increased over the years just as predicted by Moore's law.[11] The law is named after Intel cofounder Gordon E. Moore and states that the amount of computing power that can be purchased for the same amount of money doubles about every two years. This law has proven correct. We have seen computers becoming faster and memory more abundant. Similarly, storage space has expanded through *cloud computing*, which refers to the ability to access highly scalable computing resources through the Internet. Cloud computing is often available at a lower cost than that required for installation on in-house computers. This is because resources are shared across many users. Further, the performance of the algorithms that drive so many of our systems has also increased. Therefore, the gains from big data are a combination of the size of current data sets coupled with rapidly increasing processing capability and improved algorithms.

Hadoop is an open source technology platform that has received a great deal of buzz as it was designed to solve problems with lots of data. In fact, Hadoop was specifically designed to deal with big data that is a mixture of complex and structured data that does not fit nicely into tables. It was originated by Google for its own use for indexing the Web and examining user behavior to improve performance algorithms. Yahoo! then furthered its development for enterprise purposes. Hadoop uses distributed applications across many servers. The database is distributed over a large number of machines. Spreading data over multiple machines greatly improves computing capability. Because the tables are divided and distributed into multiple servers, the total number of rows in each table in each database is reduced. This reduces index size and substantially improves search performance. The database is typically divided into partitions called *database shards.* A database shard can be placed on separate hardware and multiple shards can be placed on multiple machines. Database shards significantly improve performance. The segment of the database placed on a shard can be based on real-world segmentation. This can greatly help analysis—such as separating Canadian customers versus American customers. This makes it especially easy to query a particular segment of the data or evaluate comparisons across segments.

1.1.5 New Problem Solving

Just a few years ago, the topic of analytics and computing power would have concerned only a few data geeks. Today, however, big data is an imperative for all business leaders across every industry and sector—from health care to manufacturing. The ability to capture, store, aggregate, and combine data—and then perform deep analyses—has now become accessible to virtually all organizations. This will continue as costs of computing power, digital storage, and cloud computing continue to drop. These advancements will further break down technology barriers and level the playing field, especially for small and medium-sized firms. Just consider that today an individual can purchase a disk drive with the capacity to store all of the world's music for less than $600.[12] In fact, the cost of storing a terabyte of data has fallen from $1 million in the 1970s to $50 today.[13]

What does this mean for business? Simply put, increasingly sophisticated analytical techniques combined with growing computer horsepower mean extracting unparalleled business insights. It means new and revolutionary problem-solving capability. Big data is not about the data itself. It is about the ability to solve problems better than ever before. And now pretty much everyone can do it.

1.2 What Is Different?

Companies have been capturing data for years and conducting analysis to gain market intelligence. It is natural to wonder what exactly is different today.

The difference is *scale*. This scale is in terms of the amount of data *and* the computing capability to analyze it. Combined, these elements can offer objective, evidence-based insights into virtually every aspect of a business. This has created a new level of competitiveness and the opportunity for companies to use this intelligence as a competitive advantage. We are living through a data explosion. According to a recent *New York Times* article, "Data is a vital raw material of the information economy, much as coal and iron ore were in the Industrial Revolution. But the business world is just beginning to learn how to process it all."[14]

The three characteristics of big data—volume, velocity, and variety—are what make big data different from all the other data collected in the past (see Figure 1.2).[15]

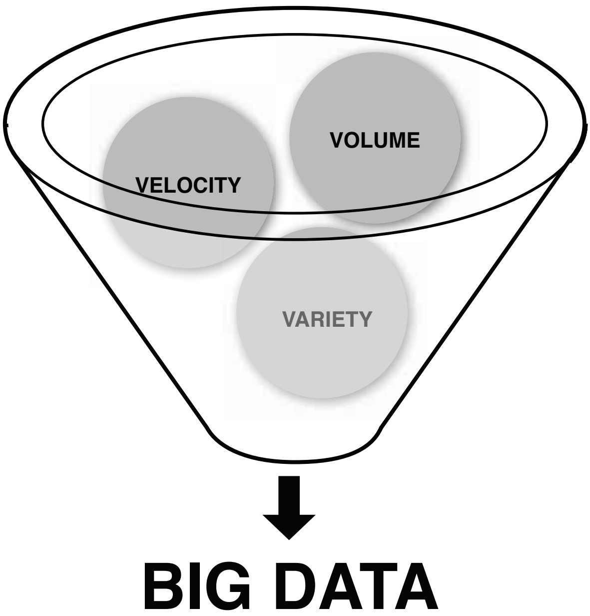

Figure 1.2 The three Vs of big data

1.2.1 Volume

Today's data is huge and data is everywhere. Consider that Google receives more than three billion queries every day, a volume that is thousands of times the quantity of all printed material in the U.S. Library of Congress; Facebook gets more than 10 million new photos uploaded every hour; Walmart conducts a million transactions per hour; and the New York Stock Exchange (NYSE) trades 490 million shares per day.[16] The amount of data generated will continue to grow exponentially. In fact, the number of RFID tags sold globally is projected to rise from 12 million in 2011 to 209 billion in 2021.[17]

Manufacturers and retailers are collecting data all along their supply chains. This includes data from POS, GPS, and RFID data, to data emitted by equipment sensors, to social media feeds. Virtually all companies have information technology (IT) systems in their back offices. The world we live in is enveloped in data. When converting terabytes and exabytes into meaningful terms, it is estimated that the data that companies and individuals are producing and storing is equivalent to filling more than 60,000 U.S. Libraries of Congress.[18] Where is this data going? It is accumulating in large pools growing larger by the minute. This is *big data.* In fact, many companies do not even recognize the data they possess and how valuable it is. Data can be traded or sold, and it has economic value. Data is the new asset.

1.2.2 Velocity

Yes, the data is large. The data is growing. And it is growing rapidly.

Data has become a deluge flowing into every aspect of business and everyday life. Companies are capturing exponentially growing volumes of transactional data. They are also capturing volumes of information about their customers, suppliers, and operations. Consider that there are millions of sensors embedded in physical devices all around us, including mobile phones, smart energy meters, automobiles, and industrial machines. These devices capture and communicate data in what is called the age of the *Internet of Things*.[19]

Companies and individuals are generating a tremendous amount of digital *exhaust data.* This is data that is created as a by-product of other activities and is generated just by going about everyday business. Consumers going about their day through texting, communicating, browsing, searching, and buying are creating *digital trails.* These trails can be captured, monitored, and analyzed. Social media sites, smartphones, PCs, and laptops have enabled billions of individuals around the world to continue adding to the amount of big data available. The growing amount of multimedia content has played a major role in the exponential growth in the amount of big data. In fact, each second of high-definition video, for example, generates more than 2,000 times as many bytes as required to store a single page of text.[20]

Companies are collecting data with increasingly greater granularity and frequency. They can capture every customer transaction. They can attach increasingly personal information to these transactions. As a result, they are collecting more information about consumer behavior in many different environments. Companies can then aggregate and disaggregate this data in infinite combinations to mathematically optimize and determine the best marketing strategies, custom-tailored shopping experiences, flexible and targeted product designs, and an optimized fulfillment system.

1.2.3 Variety

Big data is also in every form imaginable. Most of us think of data as numbers neatly stored in columns and rows. However, big data is in combinations of forms and this variety is growing. It is in the form of structured data that we are familiar with but may also be readings from sensors, GPS signals from cell phones, and POS and RFID data. It is also unstructured, such as text and voice messages, social networks feeds, and blogs.

Sources of big data are everywhere—from sensors such as RFID and POS data at retail checkouts, to geolocation information transmitted from GPS signals, to vibration and heat sensors on equipment, to social media feeds (think of Twitter "re-tweets" and Facebook "likes"), to maintenance logs, to customer complaints (see Figure 1.3). Most smartphones have GPS capabilities that enable tracking. Cars alone are stuffed with chips, sensors, and software that upload performance data to the carmakers' computers when the vehicle is serviced. Typical mid-tier vehicles now have some 40 microprocessors and electronics that account for one third of their cost.[21] How else does your car manufacturer inform you that it is time for service?

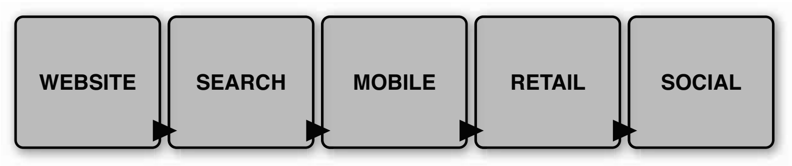

Figure 1.3 Sources of data are everywhere

1.3 What Does It Mean?

The bottom line is that big data analytics enables converting information into an unprecedented amount of business intelligence. It allows companies to precisely understand *what* happened in the past and *why*, and better predict the future. The result is a superior competitive capability.

1.3.1 Big Business Intelligence

Not using big data analytics in today's business world is akin to making deliveries on horseback while competitors are using a truck. Or using carrier pigeons while everyone else is using airfreight. A company just can't compete without it.

What can big data analytics tell us?

The size of data makes it possible to spot connections and details in the data that are otherwise impossible to spot. Granular analysis of subcategories and submarkets is enabling the understanding of customers and markets unlike ever before. A task that used to be accomplished based on traditional marketing tools—such as focus groups and surveys—to try to determine what the customer wants is now computed based on scientific methods. We no longer guess or use hunches. We know exactly what each customer wants to buy. This changes the way we sell products. This information also drives entire supply chains.

The large amounts of data enable the establishments of *norms* in the data. This also helps identify data that is outside of the norms—namely outliers. This technique, for example, is what we see today in the identification of credit card fraud. The algorithm automatically detects a change from the norm. This is only possible when there is a large amount of data. This has huge applications any time we are looking for a deviation. This is what UPS does, for example, when it monitors its vehicle fleets for preventive maintenance.[22] The company cannot afford a breakdown in its fleet of vehicles. Before big data analytics, the company routinely replaced parts, but that was wasteful. Then in 2000, the company embarked on a program of using sensors to capture data on vehicle performance and notice deviations that indicate a time to intervene before a breakdown.

What else can big data analytics do?

1.3.2 Predicting the Future

Predictive analytics uses a variety of techniques—such as statistics, modeling, and data mining—to analyze current and historical facts to make predictions about the future.[23] This is one of the most significant aspects of big data analytics. It is the ability to foresee events before they happen by sensing small changes over time. For example, IBM's Watson computer uses an algorithm to predict best medical treatments[24] and UPS uses analytics to predict vehicle breakdown.[25] By placing sensors on machinery, motors, or infrastructure like bridges, we can monitor the data patterns they give off, such as heat, vibration, stress, and sound. These sensors can detect changes that may indicate looming problems ahead—essentially forecasting a problem.

Things do not break down all at once. There is a gradual wear and tear over time. In the past, our technology, sensors, and analytics were not sophisticated enough to detect these changes. Today, armed with sensor data, correlation analysis, and similar methods, we can identify the specific patterns that typically crop up before something breaks. This may be the sound of a motor, excessive heat from an engine, or vibration from the bridge. In health care, it may be changes in a patient's vitals before the onset of disease. Google is famous for identifying location and propagation of the flu by simply tracking the volume and type of queries in its search engine.[26]

1.3.3 Fewer Black Swans

Black swans is a term used to describe high-impact, low-probability events.[27] Historically, we assumed that these could not be predicted. However, with big data analytics, that is rapidly changing. With big data analytics, the number of events that we used to consider unpredictable and purely random is getting smaller. We are now able to identify and spot changes in systems that indicate potential failure. Just consider the accuracy of the prediction of hurricane Sandy provided by the NOAA weather satellite. Only a few years earlier, this type of event would have been considered a black swan.

Spotting the abnormality early on enables the system to send out a warning so that a new part can be installed, preparation before an impending tsunami can be made, or the malfunction can be fixed. The aim is to identify and then watch a good proxy for the event we are trying to forecast, and thereby predict the future. External events—such as weather, traffic, or road construction—can be tracked and the supply chain can respond. Traffic can be rerouted or knowledge of outbreaks of flu can be used to determine which areas may need more supplies of ibuprofen, chicken soup, or cough drops. This ability is a game changer for risk management.

1.3.4 Explain What *Has Happened*

One of the most powerful analytics tools we can use on big data is *correlation analysis.* Correlation analysis has been used for decades. What is different today are the insights obtained when applied to the huge amounts of data. Correlations tell us whether there is a relationship between any set of variables. It doesn't tell us why there is a relationship. In the world of statistics, it is "quick and dirty" but offers important insights.

Correlations let us analyze a phenomenon by identifying a useful proxy for it. The idea is that if A often takes place together with B, we need to monitor B to predict that A will happen.

Consider the case of Target and identification of pregnant customers.[28] Big data analytics was able to identify the precise purchasing bundle associated with a female customer in the second trimester of pregnancy. Those who have seen the highly publicized story might recall the father of a 16-year-old girl who was very angry at Target for sending his daughter baby coupons—only to discover that indeed she was pregnant. The analytics perfectly targeted her—no pun intended.

Correlation analysis also points the way for causal investigations, by telling us which two things are potentially connected. This then tells us where to investigate further. This provides information on where to go into modeling, causation, and optimization. This is an important benefit. It points us in the right direction and enables us to know where to dig deep with more sophisticated analytics applications such as supply chain optimization.

1.3.5 Explain Why *Things Happen*

Correlation analysis tells whether a relationship exists. More advanced statistical applications enable us to go beyond understanding whether a relationship exists and delve deeper into understanding causations.

1.3.5a Supply Chain Optimization

Supply chain optimization is the application of mathematical and statistical tools to develop optimal solutions to supply chain problems. This enables analysts to create models to simulate, explore contingencies, and optimize supply chains. Many of these approaches employ some form of linear programming software and solvers. This allows the program to maximize a particular goal given a set of variables and constraints. This includes the optimal placement of inventory within the supply chain, minimizing the carbon footprint or minimizing operating costs, such as manufacturing, transportation, and distribution costs.

1.3.5b Randomized Testing

Big data and analytics have enabled companies to use randomized testing to conduct experiments to "test and learn"—sometimes called *design of experiments*. Randomized testing is a statistical method that involves conducting, analyzing, and interpreting tests to evaluate which factors impact variables of interest. For example, this might be asking whether planned changes in delivery or store layouts will increase customer purchases. Randomized testing is at the heart of the scientific method. Without random assignment to test groups, and without a control group, it is impossible to know which improvements are actually due to the changes being made. This type of large-scale testing is now possible as there is lots of data to compare and analyze.

Another significant enabler is that many current software applications are designed for people with little statistical training. New software makes it possible to conduct design of experiments by businesspeople rather than professional statisticians. For example, testing alternative versions of Web sites is relatively straightforward. This type of testing is simple and is becoming widely practiced in online retailers. Simple A/B experiments, such as comparing two versions of a Web site—A versus B—can be easily structured. The online retailer eBay, for example, routinely conducts experiments with different aspects of its Web site.[29] The site generates huge amounts of data as there are more than a billion page views per day. This enables eBay to conduct multiple experiments concurrently and not run out of treatment and control groups. Similarly, the North Carolina food retailer Food Lion uses testing to try out new retailing approaches—again simply comparing A versus B.[30] This ranges from comparing new store formats to simple tactical decisions.

1.4 Transformations

1.4.1 Business Ramifications

Consider the following examples of companies that have implemented big data analytics:

- The global cement giant CEMEX has successfully applied analytics to its distinctive capability of optimized supply chains and delivery times.[31]
- Walmart relies extensively on analytics to run its entire supply chain.
- At Deere & Company, a new way of optimizing inventory saved the company $1.2 billion in inventory costs between 2000 and 2005.[32]
- Proctor & Gamble used operations research methods to reorganize sourcing and distribution approaches in the mid-1990s and saved the company $200 million in costs.[33]
- Amazon claims its latest advanced analytics can now predict purchases *before* they happen. Based on the pattern of customer computer searches and how long the cursor lingers over a Web site, the company plans to start bundling and shipping items to distribution centers in advance of actual purchases.[34]

Questions that were once based on intuition and guesswork can now be answered in objective and quantifiable terms. Big data analytics answers business questions such as the following:

- What does the future look like? What do our customers want?
- What is the reason for our success? Is our strategy working?
- What activities should we pursue in the future? Which resources should we invest in?
- What do we do to minimize our risk exposure? How do we protect ourselves from business disruptions?

1.4.2 Changing the Present and Future

The ability to answer these questions changes virtually every aspect of business. It enables understanding both the present and future. As such, it can enhance a company's competitive position by better predicting competition and markets. It can dramatically improve operational and supply chain performance. For companies that have implemented big data analytics, it can increase productivity and improve efficiency, quality, and preventive maintenance. It can help manage suppliers and customers, as well as logistics and transportation operations. It can also better evaluate strategy, improve forecasting, help prepare for disruptions, and, overall, improve risk management.

Harnessing big data analytics has the potential to improve efficiency and effectiveness, to enable organizations to do more with less, to produce higher-quality outputs, and to increase the value-added content of their products and services. Companies can leverage their data to design products that better match customer needs. No more guessing what the customer wants. Through in-store behavior analysis and customer microsegmentation, companies can optimize market segments and know exactly what the customer is buying. In fact, analytics is moving businesses into an era of "shopper marketing"—monitoring and creating an entire shopping experience for customers no matter where they are along their shopping path: at home (online), on the go (through geolocation), and within stores (in-store monitoring).[35]

Data can even be leveraged to improve products as they are used. An example is a mobile phone that has learned the owner's habits and preferences—that holds applications, photos, and data tailored to that particular user's needs. That device will therefore become more valuable with use than a new device that has not become customized.[36]

1.4.3 Creating New Business Opportunities

The information potential of data is opening all kinds of new business opportunities. Consider the possibilities from the mere ability to gather data about how car parts are actually used on the road. This data can be used to improve the design of parts and is turning out to be a big competitive advantage for the firms that can obtain the information. Consider the company Intrix, which collects geolocation information. In 2012, the company ran a trial of analyzing where and when the automatic braking systems (ABS) of a car kicked in.[37] The premise was that frequent triggering of the ABS on a particular stretch of road may imply that conditions there are dangerous, and that drivers should consider alternative routes. With this, Intrix developed the service offering to recommend not only the shortest route, but the safest one as well. This is an entirely novel area of business.

Big data is also helping create entirely new types of businesses, especially those that aggregate and analyze data. Data is the new asset and most organizations are unable to keep up with its rapid growth, scale, and evolution. This is not their core competency. As a result, most non-IT companies are turning to some solutions providers for help. Companies are routinely outsourcing this capability and turning to third parties. This is the rise of the *third-party analytics provider (3PA)*. These are various analytics and IT experts, data brokers, software vendors, and solutions consultants. Similar to third-party logistics providers (3PLs) that orchestrate the movement of physical goods, these companies coordinate and make sense out of large data flows.

1.5 Data-Driven Supply Chains

Few areas of business have been transformed by big data analytics as much as supply chain management. Same-day delivery has become nearly mandatory to modern multichannel retailing.[38] As consumers, we have developed this expectation. We don't think about it unless there is a problem. It may be that the item we ordered online doesn't show up as scheduled or an advertised item is out of stock when we try to purchase it. Achieving a competitive level of global supply chain excellence cannot be accomplished without data-driven, end-to-end operations.

Consider companies such as Tesco. The company gathers huge amounts of customer data from its loyalty program. It then mines this data to inform decisions from promotions to strategic segmentation of customers. Amazon came early to the frontier of data analytics. The online retailer pushed the frontier using customer data to power its recommendation engine "you may also like..." based on a type of predictive modeling technique called collaborative filtering. The company continues its rapid leadership in fulfillment capabilities through data-driven decisions. Walmart was also an early adopter of data-driven supply chains. By making supply-and-demand signals visible between retail stores and suppliers, the company optimizes all its supply chain decisions—from customer fulfillment to inventory tracking (think POS data and RFID sensors) to automatic purchase orders through its supplier portal.

The number of RFID tags sensing inventories across supply chains is in the millions. In fact, the number of RFID tags sold globally is projected to rise from 12 million in 2011 to 209 billion in 2021.[39] Supply chains are increasingly combining data from different systems to coordinate activities across the supply chain end-to-end. Marketing is generating huge volumes of POS data from retail stores that is automatically shared with suppliers for real-time, stock-level monitoring. RFID tags monitor inventory on shelves and in-transit coordinating with current stock levels for automatic order replenishment. Add to this data from computer-aided design, computer-aided engineering, computer-aided manufacturing, collaborative product development management, and digital manufacturing, and connect it across organizational boundaries in an end-to-end supply chain.

Even more value can be unlocked from big data when companies are able to integrate data from other sources. This includes data from retailers that goes well beyond sales. It may be promotion data, such as items, prices, and sales. It also includes launch data, such as specific items to be listed and associated ramp-up and ramp-down plans. It also includes inventory data, such as stock levels per warehouse and sales per store. This data is essential for the supply chain to deliver the items that are needed when they are needed.

Through collaborative supply chain management and planning, companies can mitigate the bullwhip effect and better smooth out flow through the supply chain. Many companies guard customer data as proprietary, but there are many examples of successful data sharing. Walmart is a great example of requiring all suppliers to use its Retail Link platform.[40] The exchange and sharing of data across the extended enterprise has provided transparency and enabled coordinated cross-enterprise efforts.

Big data analytics is the game changer. It has given rise to the *intelligent supply chain*.

Index

N

O

P

T

U

V

W-X-Y-Z